

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 2, 2004, 19:47:31 ; Search time 5.78967 Seconds

(without alignments)
371,762 Million cell updates/sec

Title: US-10-054-873-3

Perfect score: 33

Sequence: 1 LGTPR 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	6	2	AAV42857
2	33	100.0	15	7	ADM08409
3	33	100.0	15	7	ADM08322
4	33	100.0	107	2	AAV42860
5	33	100.0	135	4	AAAB93957
6	33	100.0	135	7	ADJ68719
7	33	100.0	150	2	AAV42861
8	33	100.0	388	7	AAO69171
9	33	100.0	392	4	AAW42128
10	33	100.0	405	4	ABG22253
11	33	100.0	585	4	AAW40342
12	33	100.0	952	7	ADM04599
13	33	100.0	1547	6	ABU96680
14	33	100.0	1562	5	ABP43703
15	33	100.0	2013	5	ABBA47751
16	33	100.0	2013	6	ABU32684
17	33	100.0	7257	6	AAV58576
18	31	99.9	238	4	AAU27512
19	31	99.9	347	8	ADO29145
20	31	99.9	448	5	ABP66201
21	30	90.9	24	4	ABB39575
22	30	90.9	24	4	AAW31334
23	30	90.9	24	4	ABB24291
24	30	90.9	24	4	AAW72803
25	30	90.9	24	4	AAW60274

26	30	90.9	24	4	ABG54611	ABG54611 Human liv
27	30	90.9	24	5	ABG42737	ABG42737 Human pep
28	30	90.9	100	4	AAU49418	AAU49418 Protiomb
29	30	90.9	100	6	ABW45937	ABW45937 Protiomb
30	30	90.9	108	3	AAW04045	AAW04045 Human sec
31	30	90.9	110	4	AAO02964	AAO02964 Human sec
32	30	90.9	114	4	AAW89914	AAW89914 C Glytam1
33	30	90.9	203	4	AAW79780	AAW79780 Coryneb
34	30	90.9	209	4	AAW94054	AAW94054 Human pro
35	30	90.9	222	6	ADA03124	ADA03124 Nucleotid
36	30	90.9	222	4	AAU68588	AAU68588 Human nov
37	30	90.9	231	6	ABO00555	ABO00555 Novel hum
38	30	90.9	288	4	ABW67569	ABW67569 Amino aci
39	30	90.9	288	2	ABO07232	ABO07232 Human p53
40	30	90.9	298	2	AAV73880	AAV73880 Human pro
41	30	90.9	298	3	AAW43779	AAW43779 Human can
42	30	90.9	298	4	AAU23487	AAU23487 Novel hum
43	30	90.9	298	4	AAW25792	AAW25792 Human pro
44	30	90.9	301	5	ABW60645	ABW60645 Human NS
45	30	90.9	325	6	ABU96711	ABU96711 Human nuc

ALIGNMENTS

RESULT 1
AAV42857
ID AAV42857 standard; peptide; 6 AA.

AC AAV42857;
DT 19-JAN-2000 (first entry)

DE Cleavable peptide linker for hGH-mini-proinsulin chimeric protein.

DE Linker: growth hormone; chaperone; intramolecular; insulin; precursor;

KM folding; conformation; chimeric protein; cleavable; recombinant;

KM production; yield.

OS Synthetic.

PN WO950302-A1.

PR 07-OCT-1999.

PF 31-MAR-1998; 98WO-CN000052.

PR 31-MAR-1998; 98WO-CN000052.

(TONG-) TONGHUA GAMTECH BIOTECHNOLOGY LTD.

Gan Z;
WPI; 1999-610839/52.

New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.

Claim 6; Page 29; 46pp; English.

This sequence represents a cleavable peptide linker which is a component of the chimeric proteins hGH-mini-proinsulin (AAV42860) and the chimeric protein given in AAV42861. These chimeric proteins additionally contain an N-terminal fragment of human growth hormone (hGH) and a human insulin precursor (AAV42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue (AAV42857) which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular

degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein, and decrease the intermolecular interactions among the fusion proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative sulphydrololysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins

Sequence 6 AA;

Query Match 100.0%; Score 33; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1 LGTGR 6

RESULT 2
ADM08409 standard; peptide; 15 AA.

ADM08409;

20-MAY-2004 (first entry)

Canine immunoglobulin group 3 VL species framework 2 peptide 22.

canine; dog; heavy; immunoglobulin; antibody light chain variable domain; antiallergic; allergy; IGE; gene therapy; group 3 species; VL framework; FR2.

Canis familiaris.

WO2003060080-A2.

24-JUL-2003.

20-DEC-2002; 2002WO-US041362.

21-DEC-2001; 2001US-0344874P.

(IDEX-) IDEXX LAB INC.

Krah ER, Guo H, Aiyappa A, Lawton R;

WPI; 2003-598521/56.

New canine heavy and light chain variable domain polypeptides, useful for treating canine allergy.

Claim 40; Page 107; 130pp; English.

The invention relates to a novel canine heavy or light chain variable domain polypeptide. The protein of the invention demonstrates antiallergic activity and may be useful for treating canine allergy, possibly via gene therapy. The current sequence is that of a canine immunoglobulin light chain variable domain framework (FR) peptide of the invention.

Sequence 15 AA;

Query Match 100.0%; Score 33; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 6 LGTGR 11

RESULT 3
ADM08322 standard; peptide; 15 AA.

ADM08322;

20-MAY-2004 (first entry)

Canine immunoglobulin group 3 VL subgenus framework 2 peptide 8.

canine; dog; heavy; immunoglobulin; antibody light chain variable domain; antiallergic; allergy; IGE; gene therapy; group 3 subgenus; VL framework; FR2.

Canis familiaris.

WO2003060080-A2.

24-JUL-2003.

20-DEC-2002; 2002WO-US041362.

21-DEC-2001; 2001US-0344874P.

(IDEX-) IDEXX LAB INC.

Krah ER, Guo H, Aiyappa A, Lawton R;

WPI; 2003-598521/56.

New canine heavy and light chain variable domain polypeptides, useful for treating canine allergy.

Claim 39; Page 106; 130pp; English.

The invention relates to a novel canine heavy or light chain variable domain polypeptide. The protein of the invention demonstrates antiallergic activity and may be useful for treating canine allergy, possibly via gene therapy. The current sequence is that of a canine immunoglobulin light chain variable domain framework (FR) peptide of the invention.

Sequence 15 AA;

Query Match 100.0%; Score 33; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 6 LGTGR 11

RESULT 4
AAV42860 standard; protein; 107 AA.

AAV42860;

19-JAN-2000 (first entry)

hGH-mini-proinsulin chimeric protein.

Insulin; precursor; growth hormone; chaperone; intramolecular; folding; conformation; chimeric protein; cleavable; recombinant; production; yield.

Synthetic.

Homo sapiens.

WO950302-A1.

07-OCT-1999.

XX PF 31-MAR-1998; 98WO-CN000052.
 XX PR 31-MAR-1998; 98WO-CN000052.
 XX PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
 XX PI Gan Z;
 XX DR WPI; 1999-610839/52.
 XX PT New chimeric proteins containing human growth hormone fragment, used
 PT particularly for the production of human insulin.
 XX PS Claim 13; Page 30; 46pp; English.

CC This sequence represents a chimeric protein, hGH-mini-proinsulin. This
 CC chimeric protein contains an N-terminal fragment of human growth hormone
 CC (hGH) of the sequence given in AA42855, a cleavable peptide linker
 CC (AA42857), and a human insulin precursor comprising insulin A and B
 CC chains (AA42859). The hGH portion of the chimeric protein acts as an
 CC intramolecular chaperone (IMC) for the insulin precursor, enabling it to
 CC fold correctly. The cleavable peptide linker has a C-terminal Arg residue
 CC which enables the hGH portion of the chimeric protein to be removed after
 CC folding has taken place. Production of recombinant human insulin via an
 CC hGH-proinsulin chimeric protein can provide human insulin with correctly
 CC linked cysteine bridges with fewer necessary procedural steps, and hence
 CC resulting in a higher yield of human insulin. The IMC sequences not only
 CC protect insulin sequences from intracellular degradation by a
 CC microorganism host, but also promote the folding of the fused insulin
 CC precursor, facilitate the solubility of the fusion protein and decrease
 CC the intermolecular interactions among the fusion proteins, thus allowing
 CC folding of the fused insulin precursor at commercially useful high
 CC concentrations. The procedural steps of cyanogen bromide cleavage,
 CC oxidative sulphyloysis and related purification steps can thus be
 CC eliminated, along with the use of high concentrations of mercaptan or the
 CC use of hydrophobic absorbent resins

XX SQ Sequence 107 AA;

Query Match 100.0%; Score 33; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 |||||
 Db 50 LGTGR 55

RESULT 5
 AAB93957
 ID AAB93957 standard; protein; 135 AA.
 XX AAB93957;
 AC AAB93957;
 XX 26-JUN-2001 (first entry)
 DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:14002.
 DE Human protein sequence SEQ ID NO:14002.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 OS Homo sapiens.
 XX Homo sapiens.
 PN BP1074617-A2.
 XX BP1074617-A2.
 PD 07-FEB-2001.
 XX 07-FEB-2001.
 PF 28-JUL-2000; 2000EP-00116126.
 XX 28-JUL-2000; 2000EP-00116126.
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00185767.

PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX DR WPI; 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

XX PS Claim 8; SEQ ID NO 14002; 2537pp + Sequence listing; English.

CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

XX SQ Sequence 135 AA;

Query Match 100.0%; Score 33; DB 4; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 |||||
 Db 5 LGTGR 10

RESULT 6
 ADJ69719
 ID ADJ69719 standard; protein; 135 AA.
 XX ADJ69719;
 AC ADJ69719;
 XX 06-MAY-2004 (first entry)
 DT 06-MAY-2004 (first entry)
 XX Human heat mitochondrial protein as a therapeutic target SeqID:1525.
 DE Human heat mitochondrial protein as a therapeutic target SeqID:1525.
 XX Mitochondrial; human; screening assay; diabetes mellitus;
 KM Huntington's disease; osteoarthritis;
 KM Leber's hereditary optic neuropathy; LHON;
 KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KM myoclonic epilepsy ragged red fibre syndrome; MERRF; Cancer;
 KM neuroprotective; nootropic; antidabetic; anticonvulsant; antitachycardic;
 KM osteopathic; ophthalmological; cytosstatic.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO2003087768-A2.
 XX WO2003087768-A2.
 PD 23-OCT-2003.

XX	04-APR-2003; 2003WO-US010870.
XX	12-APR-2002; 2002US-0372843P.
PR	17-JUN-2002; 2002US-0389867P.
PR	20-SEP-2002; 2002US-0412418P.
XX	(MITO-) MITOKOR.
PA	(BUCK-) BUCK INST AGE RES.
XX	Gnosh SS, Fahy ED, Zhang B, Gibson SW, Taylor SW, Glenn GW;
PI	Warnock DE;
XX	WPI; 2003-845369/78.
XX	
PT	Identifying a mitochondrial target for drug screening assays and for
PT	treating diseases associated with altered mitochondrial function,
PT	comprises detecting a modified polypeptide in a sample and correlating
PT	with the disease.
XX	
PS	Claim 1; SEQ ID NO 1525; 180dp; English.
XX	
CC	This invention relates to novel mitochondrial targets that can be used
CC	for therapeutic intervention in treating a disease associated with
CC	altered mitochondrial function. Specifically, it refers to a method for
CC	identifying proteins of the human heart mitochondrial proteome that are
CC	useful for drug screening assays, as well as therapeutic targets. The
CC	present invention describes a method for identifying such proteins that
CC	can be used in the treatment of various diseases associated with altered
CC	mitochondrial function including diabetes mellitus, Huntington's disease,
CC	osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC	encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC	ragged red fibre syndrome (MERF) or cancer. Accordingly, these
CC	compositions have neuroprotective, nootropic, antidiabetic,
CC	anticoagulant, antiarthritic, osteopathic, ophthalmological and
CC	cytoskeletal activities. This polypeptide sequence is a human heart
CC	mitochondrial protein of the invention.
XX	
SQ	Sequence 135 AA;
XX	
Query Match	100.0%; Score 33; DB 7; Length 135;
Best Local Similarity	100.0%; Pred. No. 1.3e+02;
Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Yq	1 LGTGR 6
Db	5 LGTGR 10
RESULT 7	
ID	AAV42861 standard; protein; 150 AA.
AC	AAV42861;
XX	
DT	19-JAN-2000 (first entry)
XX	
DE	Chimeric protein, SEQ ID 7.
XX	
KW	Insulin; precursor; growth hormone; chaperone; intramolecular; folding;
KW	conformation; chimeric protein; cleavable; recombinant; production;
KW	yield.
XX	
OS	Synthetic.
XX	
OS	Homo sapiens.
XX	
PN	MO9950302-AI.
XX	
PD	07-OCT-1999.
XX	
DF	31-MAR-1998; 98WO-CN000052.
XX	
XX	31-MAR-1998; 98WO-CN000052.

XX	PA	(TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.
XX	PI	Gan 2;
XX	DR	WPI, 1999-610839/52.
XX	PT	New chimeric proteins containing human growth hormone fragment, used particularly for the production of human insulin.
XX	PS	Claim 14; Page 30-31; 46pp; English.
CC	CC	This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AAY42856, a cleavable peptide linker (AAI42857), and a human insulin precursor comprising insulin A and B chains (AAV42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proteinase chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from intracellular degradation by a microorganism host, but also promote the folding of the fused insulin precursor, facilitate the solubility of the fusion protein and decrease the intermolecular interactions among the insulin proteins, thus allowing folding of the fused insulin precursor at commercially useful high concentrations. The procedural steps of cyanogen bromide cleavage, oxidative subunitolysis and related purification steps can thus be eliminated, along with the use of high concentrations of mercaptan or the use of hydrophobic absorbent resins
XX	SC	Sequence 150 AA;
QY	1 LGTGPR 6	
Db	93 LGTGPR 98	
XX	QY	100.0%; Score 33; DB 2; Length 150;
XX	Db	Best Local Similarity 100.0%; Pred. No. 1,4e+02;
XX	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
XX	RESULT 8	
XX	AA069171	AB069171 standard; protein; 388 AA.
XX	AB069171;	
XX	AC	29-JUL-2004 (first entry)
XX	DT	
XX	DR	Pseudomonas aeruginosa polypeptide #1346.
XX	KM	Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX	OS	Pseudomonas aeruginosa.
XX	PN	US6551795-B1.
XX	PD	22-APR-2003.
XX	PF	18-FEB-1999; 99US-00252991.
XX	PR	18-FEB-1998; 98US-0074788P.
XX	PR	27-JUL-1998; 98US-0094190P.
XX	PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	PI	Rubenfeld MJ, Nolling J, Delonghery C, Bush D;
XX	WPI, 2003-615309/58.	
XX	N-PSDB; ABD02742.	

XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure: SEQ ID NO 17917; 455pp; English.
 XX
 CC The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant
 CC production of *P. aeruginosa*-derived peptide or polypeptides, as target
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
 CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 CC
 XX Sequence 388 AA;
 SQ
 Query Match 100.0%; Score 33; DB 7; Length 388;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 DB 179 LGTGR 184
 RESULT 9
 ID AAM42128 standard; protein; 392 AA.
 XX
 AC AAM42128;
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 7059.
 XX
 KW Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; chemolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-UTL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00588042.
 PR 19-UTL-2000; 2000US-00620332.
 PR 03-AUG-2000; 2000US-00633450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI; 2001-442253/47.
 DR N-PSDB; AA161284.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 2; SEQ ID NO 7059; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM8642-AAM42213) with nocrotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 CC
 XX Sequence 392 AA;
 SQ
 Query Match 100.0%; Score 33; DB 4; Length 392;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 DB 17 LGTGR 22
 RESULT 10
 ID ABG22253 standard; protein; 405 AA.
 XX
 AC ABG22253;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #22244.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AA586440.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 52612; 103bp; English.

PS The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping;

CC and in recombinant production of (II). The polynucleotides are also used

CC in diagnostics as expression sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is

CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food

CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders

CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 405 AA;

SO

Query Match 100.0%; Score 33; DB 4; Length 405;

Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6

Db 87 LGTGR 92

RESULT 11

AA040342

ID AA040342 standard; protein; 585 AA.

XX

AC AA040342;

DT 22-OCT-2001 (first entry)

XX

DE Human polypeptide SEQ ID NO 3487.

XX

KW Human; nocrotrophic; immunosuppressant; cytostatic; gene therapy; cancer;

KW peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemokine; chemokine;

KW chemokine; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX

OS Homo sapiens.

PN WO200153312-A1.

XX

PD 26-JUL-2001.

XX

FE 26-DEC-2000; 2000WO-US034263.

XX

PR 23-DEC-1999; 99US-00471275.

XX

PR 21-JAN-2000; 2000US-0048725.

XX

PR 25-APR-2000; 2000US-0052317.

XX

PR 20-JUN-2000; 2000US-00598042.

XX

PR 19-JUL-2000; 2000US-00620312.

XX

PR 03-AUG-2000; 2000US-00653450.

XX

PR 14-SEP-2000; 2000US-00662191.

XX

PR 19-OCT-2000; 2000US-00693036.

XX

PR 29-NOV-2000; 2000US-00727344.

XX

PA (HVSSE-) HVSSEQ INC.

XX

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,

PI Wang J, Wang Z, Weinman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

PI Zhou P, Goodrich R, Drmanac RT;

XX

DR WPI: 2001-442253/47.

XX

DR N-PSDB; AA159498.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX

PS Example 6; SEQ ID NO 3487; 10078bp; English.

XX

XX The invention relates to human nucleic acids (AA157798-AA161369) and the

CC encoded polypeptides (AA138642-AA42213) with nocrotrophic

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: immune system suppression,

CC activation/inhibition activity, chemokine/chemokine activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemia and

CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX

SO Sequence 585 AA;

Query Match 100.0%; Score 33; DB 4; Length 585;

Best Local Similarity 100.0%; Pred. No. 5.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6

Db 5 LGTGR 10

RESULT 12

ADM04599

ID ADM04599 standard; protein; 952 AA.

XX

AC ADM04599;

DT 20-MAY-2004 (first entry)

XX

DE Human protein of the invention SEQ ID NO:3284.

XX

KW human; gene therapy; diagnostic marker; pharmaceutical.

XX

OS Homo sapiens.

XX

PN EP1347046-A1.

XX

PD 24-SEP-2003.

XX

FE 12-APR-2002; 2002EP-00008400.

XX

PR 22-MAR-2002; 2002JP-00137785.

XX

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;

XX

DR WPI: 2003-723558/69.

XX

DR N-PSDB; ADM02156.

XX

PT New polynucleotides and polypeptides are useful in gene therapy, for

PT developing a diagnostic marker or medicines for regulating their

PT expression and activity, or as a target of gene therapy.

XX PS Claim 1; SEQ ID NO 3284; 305bp; English.
 XX CC The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADW06202-ADW06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADW01316-ADW03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADW03759-ADW06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC protein sequence of the invention.
 XX SQ Sequence 952 AA;
 Query Match 100.0%; Score 33; DB 7; Length 952;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTSPR 6
 Db 634 LGTSPR 639
 RESULT 13
 ABU96680
 ID ABU96680 standard; protein; 1547 AA.
 AC ABU96680;
 XX 25-JUN-2003 (first entry)
 DT Human nucleic acid-associated protein (NMAP) #9.
 XX Human; nucleic acid-associated protein; cytosolic; antiarteriosclerotic;
 KM anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV;
 KM antiallergic; anti-inflammatory; thymometric; gene therapy;
 KM cell proliferative disorder; cancer; atherosclerosis;
 KM neurological disorder; epilepsy; Huntington's disease; stroke;
 KM immune disorder; inflammatory disorder; AIDS; allergy;
 KM developmental disorder; Hypothyroidism; Cushing's syndrome; infection;
 KM protein-protein interaction; drug-target interaction;
 KM gene expression profile.
 XX Homo sapiens.
 OS WO2003023003-A2.
 XX 20-MAR-2003.
 XX 05-SEP-2002; 2002WO-US028540.
 XX 07-SEP-2001; 2001US-0317792P.
 PR 07-SEP-2001; 2001US-0317912P.
 PR 14-SEP-2001; 2001US-0322270P.
 PR 21-SEP-2001; 2001US-0324040P.
 PR 28-SEP-2001; 2001US-0326732P.
 PR 19-OCT-2001; 2001US-0346716P.
 PR 25-JAN-2002; 2002US-0351749P.
 PR 22-FEB-2002; 2002US-0359498P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Tang YT, Jackson JT, Griffin JA, Elliott VS, Forsythe JT;
 PI Becha SD, Richardson TW, Lee EA, Sprague WM, Emerling BM;
 PI Thangavelu K, Warren BA, Tran UK, Yue H, Xu Y, Yue H, Li UX;
 PI Hafalia AD, Sanjanwala B, Margulis JP, Gorvad AE, Lee SY, Ison CH;
 PI Baughn MR, Chawla NK, Nguyen DB, Swarnakar A, Zebajjadian Y, Shah P;
 PI Thornton M, Yao MG, Khan FA, Gandhi AR, Yang J, Kadle AB;
 PI Burford N, Ramkumar J;
 XX WPI; 2003-313243/30.

DR N-PSDB; ACA98928.
 XX New human nucleic acid associated proteins (NMAP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant NMAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
 PT infections.
 XX Claim 1; Page 243-247; 345bp; English.
 XX The invention describes a novel human isolated nucleic acid-associated
 CC polypeptide (NMAP). The polypeptides and polynucleotides are useful in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of NMAP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. Hypothyroidism, Cushing's syndrome)
 CC disorders, or infections. These are also useful in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of NMAP. The NMAP or its fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide.
 CC The microarray is useful in monitoring or measuring protein-protein
 CC interactions, drug-target interactions, and gene expression profiles.
 CC This is the amino acid sequence of a novel human nucleic acid-associated
 CC protein (NMAP)
 XX SQ Sequence 1547 AA;
 Query Match 100.0%; Score 33; DB 6; Length 1547;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTSPR 6
 Db 1276 LGTSPR 1281
 RESULT 14
 ABP43703
 ID ABP43703 standard; protein; 1562 AA.
 AC ABP43703;
 XX 26-FEB-2003 (first entry)
 DT Alpha -NAC muscle specific form gp220.
 XX Neuroprotective; immunomodulator; cancer; chromosome 7q31-q32;
 KM cytosolic; anti-inflammatory; gene therapy; nutritional supplement;
 KM wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
 KM amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KM vulnery.
 XX Mus musculus.
 OS WO200231111-A2.
 XX 18-APR-2002.
 XX 11-OCT-2001; 2001WO-US027760.
 PR 12-OCT-2000; 2000US-0067527.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehman T, Dmanac RT;
 XX WPI; 2002-426278/45.
 DR N-PSDB; ABC60947.
 XX New polypeptides and their encoded proteins, useful as nutritional

PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.
 XX
 XX Claim 20; SEQ ID # 606; 357bp + Sequence Listing; English.
 CC The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC inflammatory, neuroprotective, immunomodulator, cyostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records ABP43544-
 CC ABP4389 represent polypeptides encoded by polynucleotides of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 1562 AA;
 Query Match 100.0%; Score 33; DB 5; Length 1562;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 Db 1291 LGTGR 1296
 RESULT 15
 ID ABB47751
 XX ABB47751 standard; protein; 2013 AA.
 AC ABB47751;
 XX 05-FEB-2002 (first entry)
 DT
 XX
 DE Listeria monocytogenes protein #455.
 XX
 XX Listeria monocytogenes protein #455.
 KM Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM Vitamin B12; bacterial infection; disease.
 XX
 OS Listeria monocytogenes.
 OS
 XX
 PN WO20017335-A2.
 PD 18-OCT-2001.
 PF 11-APR-2001; 2001WO-FR001118.
 XX
 PR 11-APR-2000; 2000FR-00004629.
 XX
 PA (INSP) INST PASTEUR.
 XX
 PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Feihl H, Dehoux P,
 PI Dusanguet O, Cheironi F, Nedjari H, Glaser P, Kunst F, Cossart P,
 PI Daniels U, Goebel W, Krefz T, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L,
 PI Perez-Diaz J, Baguerio F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablo B, Wehland U, Kaerst U, Entian K, Haut J,
 PI Rose M, Voss H;
 XX
 DR WPI/ 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and related
 PT polypeptides.
 XX
 PS Claim 6; SEQ ID NO 456; 192bp; French.

XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms. NOTE: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 2013 AA;
 Query Match 100.0%; Score 33; DB 5; Length 2013;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 Db 692 LGTGR 697
 Search completed: November 2, 2004, 20:11:42
 Job time : 8.78967 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:02:41 ; Search time 1.46125 Seconds

(without alignments)
272.306 Million cell updates/sec

Title: US-10-054-873-3

Perfect score: 33

Sequence: 1 LGTRPR 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PTCUS COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	388	4 US-09-252-991A-17917	Sequence 17917, A
2	33	100.0	7257	3 US-09-335-409-5	Sequence 5, Appl
3	33	100.0	7257	3 US-09-568-102-5	Sequence 5, Appl
4	33	100.0	7257	3 US-09-567-969-5	Sequence 5, Appl
5	33	100.0	7257	3 US-09-568-480-5	Sequence 5, Appl
6	33	100.0	7257	3 US-09-568-486-5	Sequence 5, Appl
7	33	100.0	7257	3 US-09-568-472-5	Sequence 5, Appl
8	33	100.0	7257	3 US-09-567-899-5	Sequence 5, Appl
9	30	90.9	108	4 US-09-513-999C-8126	Sequence 8126, Ap
10	30	90.9	124	4 US-09-621-976-6823	Sequence 6823, Ap
11	30	90.9	455	4 US-09-470-767-46061	Sequence 46061, A
12	30	90.9	542	4 US-09-252-991A-32139	Sequence 32139, A
13	30	90.9	684	4 US-09-252-991A-24802	Sequence 24802, A
14	29	87.9	17	3 US-09-461-697-28	Sequence 28, Appl
15	29	87.9	20	3 US-08-724-4668-9	Sequence 9, Appl
16	29	87.9	20	3 US-08-882-164D-9	Sequence 9, Appl
17	29	87.9	58	3 US-08-827-219-39	Sequence 39, Appl
18	29	87.9	84	4 US-09-583-110-5115	Sequence 5115, Ap
19	29	87.9	143	4 US-09-252-991A-30955	Sequence 30955, Ap
20	29	87.9	153	4 US-09-621-976-4040	Sequence 4040, Ap
21	29	87.9	168	4 US-09-252-991A-14281	Sequence 14281, A
22	29	87.9	176	4 US-09-252-991A-24281	Sequence 24281, A
23	29	87.9	198	4 US-09-248-796A-17378	Sequence 17378, A
24	29	87.9	217	4 US-09-252-991A-19838	Sequence 19838, A
25	29	87.9	248	4 US-09-252-991A-21872	Sequence 21872, A
26	29	87.9	259	4 US-09-436-434-2	Sequence 2, Appl
27	29	87.9	272	4 US-09-252-991A-17461	Sequence 17461, A

28	29	87.9	274	4 US-09-252-991A-17147	Sequence 17147, A
29	29	87.9	293	4 US-09-252-991A-18793	Sequence 18793, A
30	29	87.9	305	4 US-09-710-279-1362	Sequence 1362, Ap
31	29	87.9	313	4 US-09-252-991A-23631	Sequence 23631, A
32	29	87.9	313	4 US-09-252-991A-23810	Sequence 23810, A
33	29	87.9	345	4 US-09-107-532A-4268	Sequence 4268, Ap
34	29	87.9	357	4 US-09-252-991A-28380	Sequence 28380, A
35	29	87.9	365	4 US-09-134-000C-4369	Sequence 4369, Ap
36	29	87.9	408	4 US-09-252-991A-20095	Sequence 20095, Ap
37	29	87.9	409	4 US-09-252-991A-23414	Sequence 23414, A
38	29	87.9	412	4 US-09-355-912A-5	Sequence 5, Appl
39	29	87.9	430	4 US-09-252-991A-32661	Sequence 32661, A
40	29	87.9	460	4 US-09-198-452A-1085	Sequence 1085, Ap
41	29	87.9	474	4 US-09-252-991A-29391	Sequence 29391, A
42	29	87.9	480	4 US-09-252-991A-24889	Sequence 24889, A
43	29	87.9	507	4 US-09-248-796A-15540	Sequence 15540, A
44	29	87.9	518	4 US-09-252-991A-23604	Sequence 23604, A
45	29	87.9	550	4 US-09-252-991A-21295	Sequence 21295, A

ALIGNMENTS

```
RESULT 1
US-09-252-991A-17917
; Sequence 17917, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17917
; LENGTH: 388
; TYPE: PRN
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17917

Query Match      100.0%; Score 33; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LGTRPR 6
Db      179 LGTRPR 184

RESULT 2
US-09-335-409-5
; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schnupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOCHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
```

TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-5

Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1041 LGTGR 1046

RESULT 3
US-09-568-102-5
Sequence 5, Application US/09568102
Patent No. 6346404

GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-102-5

Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1041 LGTGR 1046

RESULT 4
US-09-567-969-5
Sequence 5, Application US/09567969
Patent No. 6355457

GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,969
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-567-969-5

Query Match 100.0%; Score 33; DB 3; Length 7257;

Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1041 LGTGR 1046

RESULT 5
US-09-568-480-5
Sequence 5, Application US/09568480
Patent No. 6355458

GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,480
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-480-5

Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1041 LGTGR 1046

RESULT 6
US-09-568-486-5
Sequence 5, Application US/09568486
Patent No. 6355459

GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,486
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 7257
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-486-5

Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||

DB 1041 LGTGR 1046

```
RESULT 7
US-09-568-472-5
; Sequence 5, Application US/09568472
; Patent No. 6358719
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Lignon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devor
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,472
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-472-5
```

```
Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LGTGR 6
DB 1041 LGTGR 1046

```
RESULT 8
US-09-567-899-5
; Sequence 5, Application US/09567899
; Patent No. 6383787
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Lignon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devor
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,899
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-567-899-5
```

```
Query Match
Best Local Similarity 100.0%; Score 33; DB 3; Length 7257;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LGTGR 6
DB 1041 LGTGR 1046RESULT 9
US-09-513-999C-8126

; Sequence 8126, Application US/09513993C

```
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclercq, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,993C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent .pm
; SEQ ID NO 8126
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 101
; OTHER INFORMATION: Xaa=ile or Met
US-09-513-999C-8126
```

```
Query Match
Best Local Similarity 83.3%; Score 30; DB 4; Length 108;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LGTGR 6
DB 36 VGTGR 41

```
RESULT 10
US-09-621-976-6823
; Sequence 6823, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent .pm
; SEQ ID NO 6823
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6823
```

```
Query Match
Best Local Similarity 90.3%; Score 30; DB 4; Length 124;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 LGTGR 6
DB 36 VGTGR 41

```
RESULT 11
US-09-270-767-46061
; Sequence 46061, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 46061
LENGTH: 455
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-46061

Query Match 90.9%; Score 30; DB 4; Length 455;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
DB 1.07 LGTGR 112

RESULT 12
US-09-252-991A-32139
Sequence 32139, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32139
LENGTH: 542
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32139

Query Match 90.9%; Score 30; DB 4; Length 542;
Best Local Similarity 83.3%; Pred. No. 6.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
DB 261 VGTGR 266

RESULT 13
US-09-252-991A-24802
Sequence 24802, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24802
LENGTH: 684
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24802

Query Match 90.9%; Score 30; DB 4; Length 684;
Best Local Similarity 83.3%; Pred. No. 8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
DB 274 VGTGR 279

RESULT 14
US-09-461-697-28

Sequence 28, Application US/09461697
Patent No. 627974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kasturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 17
TYPE: PRT
ORGANISM: Homo sapiens
US-09-461-697-28

Query Match 87.9%; Score 29; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GTGR 6
DB 11 GTGR 15

RESULT 15
US-08-724-466B-9
Sequence 9, Application US/08724466B
Patent No. 6063606
GENERAL INFORMATION:

APPLICANT: Petkovich, P. Martin, White, Jay A.
APPLICANT: Beckett, Barbara R., Jones, Glenville
TITLE OF INVENTION: Retinoid Metabolizing Protein
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
ZIP: M5L 1A9
COUNTRY: Canada
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
COMPUTER: COMPAQ, IBM PC compatible
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,466B
FILING DATE: October 1, 1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/667,546
FILING DATE: June 21, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 50767/00004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344

Sat Nov 6 18:59:22 2004

us-10-054-873-3.ra1

Page 5

TELEFAX: (416) 863-2653
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-724-466B-9

Query Match 87.9%; Score 29; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 GTGPR 6
Db 3 GTGPR 7

Search completed: November 2, 2004, 20:24:33
Job time : 2.46125 secs

This Page Blank (uspto)

Sat Nov 6 18:59:23 2004

us-10-054-873-3.rpr

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:59:41, Search time 1.10701 Seconds
(without alignments)
521.495 Million cell updates/sec

Title: US-10-054-873-3

Perfect score: 33

Sequence: 1 LGTGR 6

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: PIR.79.*
2: PIR1.*
3: PIR2.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	68	2 AF2809	hypothetical prote
2	33	100.0	931	2 T49710	related to glucan
3	33	100.0	2013	2 AD1129	probable peptidogl
4	33	100.0	2013	2 A11469	conserved hypochet
5	31	93.9	188	2 C87341	hypothetical prote
6	30	90.9	329	2 T44603	probable guanylate
7	30	90.9	387	2 C84847	guanylate kinase (
8	30	90.9	387	2 T50675	glutamate kinase-1
9	30	90.9	391	2 T46191	glutamate-5-semial
10	30	90.9	432	2 S49980	fasciclin II, tran
11	30	90.9	811	2 A41054	fasciclin II, PI-11
12	30	90.9	873	2 A41054	Ca2+-transporting
13	30	90.9	1002	2 A36691	hypothetical prote
14	29	87.9	85	2 G70824	conserved hypochet
15	29	87.9	90	2 B95105	hypothetical prote
16	29	87.9	100	2 B97973	hypothetical prote
17	29	87.9	136	2 T36624	hypothetical prote
18	29	87.9	169	2 A84330	hypothetical prote
19	29	87.9	212	2 T44591	hypothetical prote
20	29	87.9	219	1 TUBP2	tail fiber protein
21	29	87.9	226	2 S27759	maturation-associat
22	29	87.9	232	1 NMBR13	25.5K membrane pro
23	29	87.9	238	2 T40820	proline-rich prote
24	29	87.9	243	2 S27758	maturation-associat
25	29	87.9	243	2 A33263	expd protein limpo
26	29	87.9	256	2 T11669	protein CPD27 dr
27	29	87.9	259	2 T37915	hypothetical prote
28	29	87.9	260	2 A36949	28.5K basic DNA-bi
29	29	87.9	302	2 T15936	hypothetical prote

30	29	87.9	325	2 T35271	probable transcrip
31	29	87.9	359	2 T35179	vanillate O-demeth
32	29	87.9	371	1 HUBP1A	hyaluronoglucosam
33	29	87.9	371	2 B39625	T-cell receptor al
34	29	87.9	397	2 A39565	lymphoid enhancer-
35	29	87.9	399	2 A39625	T-cell receptor al
36	29	87.9	412	1 A42924	[3-methyl-2-oxobut
37	29	87.9	412	2 C72548	probable dihydrotri
38	29	87.9	460	2 A72009	fumarate hydratase
39	29	87.9	460	2 B86617	fumarate hydratase
40	29	87.9	461	2 E71672	fumarate hydratase
41	29	87.9	463	2 B81725	fumarate hydratase
42	29	87.9	463	2 D97826	fumarate hydratase
43	29	87.9	463	2 H71462	probable fumarate
44	29	87.9	463	2 D87510	fumarate hydratase
45	29	87.9	464	2 H83538	fumarate hydratase

ALIGNMENTS

RESULT 1
AF2809
hypothetical protein Atut1896 [imported] - Agrobacterium tumefaciens (strain C58, Dupon
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #ext_change 09-Jul-2004
C/Accession: AF2809
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guentherer, D.; Kutyavyn, T.; Levy, R.; Li, M.; McCle
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; NUID:1608550; PMID:11743193
A/Accession: AF2809
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-68 <KOR>
A/Cross-references: UNIPROT:Q8UE65; GB:AE008688; PIDN:AAL42892.1; PID:G17740345; GSPDB
A/Experimental source: strain C58 (Dupont)
A/Genetics:
A/Gene: Atut1896
A/Map position: circular chromosome

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 LGTGR 6
DB 36 LGTGR 41

RESULT 2
T49710
related to glucan 1,4-alpha-glucosidase [imported] - Neurospora crassa
N/Alternate names: protein B23121.230
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #ext_change 09-Jul-2004
C/Accession: T49710
R/Schulze, U.; Rign, V.; Hehnel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49710
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-931 <SCH>
A/Cross-references: UNIPROT:Q9P5K6; EMBL:AL356172; GSPDB:GN00116; NCSP:B23121.230
A/Experimental source: BAC clone B23121; strain OR74A
A/Genetics:
A/Gene: NCSP:B23121.230
A/Map position: 6

A:introns: 503/2

Query Match 100.0%; Score 33; DB 2; Length 931;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 797 LGTGR 802

RESULT 3

probable peptidoglycan bound protein (LPXTG motif) lmo0435 [imported] - *Listeria monocytogenes*
 C:Species: *Listeria monocytogenes*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AD1129
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entlian, K.D.; Fahl, H.; D.; Jones, L.M.; Karet, U.
 Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kuster, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schuener, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of *Listeria* species
 A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AD1129
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2013 <GUA>

A:Cross-references: UNIPROT:Q8Y9T8; GB:NC_003210; PIDN:CAC98514.1; PID:G16409812; GSPDB:
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo0435

Query Match 100.0%; Score 33; DB 2; Length 2013;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 892 LGTGR 897

RESULT 4

probable peptidoglycan bound protein (LPXTG motif) lln0457 [imported] - *Listeria innocua*
 C:Species: *Listeria innocua*
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AT1489
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
 A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entlian, K.D.; Fahl, H.
 D.; Jones, L.M.; Karet, U.
 Science 294, 849-852, 2001

A:Authors: Kretz, J.; Kuhn, M.; Kuster, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schuener, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
 A:Title: Comparative genomics of *Listeria* species
 A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AT1489
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2013 <GUA>

A:Cross-references: UNIPROT:Q92EK2; GB:AL592022; PIDN:CAC95689.1; PID:G16412898; GSPDB:
 A:Experimental source: strain C13D11262
 C:Genetics:
 A:Gene: lln0457

Query Match 100.0%; Score 33; DB 2; Length 2013;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 892 LGTGR 897

RESULT 5

conserved hypothetical protein CC0742 [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C/Accession: C87341
 R/Nierman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87341
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-188 <STO>

A:Cross-references: UNIPROT:Q9AA64; GB:AE005673; MID:G13421975; PIDN:AAK22727.1; GSPDB:
 C:Genetics:
 A:Gene: CC0742

Query Match 93.9%; Score 31; DB 2; Length 188;
 Best Local Similarity 83.3%; Pred. No. 43;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 130 MGTGR 135

RESULT 6

hypothetical protein CG1-83 [imported] - human
 C:Species: *Homo sapiens* (man)
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C/Accession: T44603
 R/Lin, W.C.
 submitted to the EMBL Data Library, May 1999

A:Description: Comparative gene cloning: Identification of novel human genes with Caeno
 A:Reference number: 222808
 A:Accession: T44603
 A:Status: preliminary; translated from GB/EMBL/DBS

A:Molecule type: mRNA
 A:Residues: 1-288 <LIN>
 A:Cross-references: UNIPROT:Q9Y392; EMBL:AF151841; PIDN:AAD34078.1

A:Genetics:
 A:Map position: 8

Query Match 90.9%; Score 30; DB 2; Length 288;
 Best Local Similarity 83.3%; Pred. No. 1.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 36 VGTGR 41

RESULT 7

probable guanylate kinase [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C/Accession: C84847
 R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentel, M.L.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umeyama, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: C84847
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-329 <STO>
 A:Cross-references: UNIPROT:P93757; GB:AE002093; NID:g6598818; PIDN:AAF18683.1; GSPDB:GN
 C:Genetics:
 A:Gene: AT2941880
 A:Map position: 2

Query Match 90.9%; Score 30; DB 2; Length 329;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 71 LGTGR 76

RESULT 8

T50675
 guanylate kinase (EC 2.7.4.8) [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
 C:Accession: T50675
 R:Kumar, V.; Spangenberg, O.; Konrad, M.
 Eur. J. Biochem. 267, 606-615, 2000
 A:Title: Cloning of the guanylate kinase homologues AGK-1 and AGK-2 from Arabidopsis th
 A:Reference number: Z25173; PMID:20098538; PMID:10632732
 A:Accession: T50675
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-387 <KUM>
 A:Cross-references: UNIPROT:Q9M681; EMBL:AF204677; PIDN:AAF60252.1
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Gene: AGK-1
 A:Introns: 1/3; 39/3; 65/2; 108/3; 229/3; 315/3; 331/1; 361/2
 C:Keywords: phosphotransferase

Query Match 90.9%; Score 30; DB 2; Length 387;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 71 LGTGR 76

RESULT 9

T46191
 guanylate kinase-like protein - Arabidopsis thaliana
 N/Alternate names: protein T8H10.150
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 C:Accession: T46191
 R:Benne, V.; Reichmann, S.; Bokkova, D.; Ansoorge, W.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: Z23014
 A:Accession: T46191
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-391 <BEN>
 A:Cross-references: UNIPROT:Q9SCB8; EMBL:AL133248
 A:Experimental source: cultivar Columbia; BAC clone T8H10
 C:Genetics:
 A:Map position: 3
 A:Introns: 4/1; 40/3; 66/2; 109/3; 230/3; 271/2; 316/3; 332/1; 364/2
 A:Note: T8H10.150

Query Match 90.9%; Score 30; DB 2; Length 391;
 Best Local Similarity 83.3%; Pred. No. 1.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 72 LGTGR 77

RESULT 10

S49980
 glutamate-5-semialdehyde dehydrogenase (EC 1.2.1.41) - Corynebacterium glutamicum (fr.
 N/Alternate names: gamma-glutamyl phosphate reductase
 C:Species: Corynebacterium glutamicum
 C>Date: 13-Jan-1995 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S49980
 R:Seberjnski, I.; Wojcik, F.; Reyes, O.; Leblon, G.
 submitted to the EMBL data library, November 1994
 A:Description: Two loci of Corynebacterium glutamicum ATCC17965 that complement Esche
 A:Reference number: S49977
 A:Accession: S49980
 A:Molecule type: DNA
 A:Residues: 1-432 <SER>
 A:Cross-references: UNIPROT:P45638; EMBL:X82929; NID:g599719; PIDN:CAA58103.1; PID:g5
 C:Genetics:
 A:Gene: proA
 C:Superfamily: gamma-glutamyl phosphate reductase
 C:Keywords: oxidoreductase

Query Match 90.9%; Score 30; DB 2; Length 432;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 36 LGTGR 41

RESULT 11

A41054
 fasciclin II, transmembrane splice form precursor - fruit fly (Drosophila melanogaster
 C:Species: Drosophila melanogaster
 C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 09-Jul-2004
 C:Accession: A41054
 R:Grenningloh, G.; Rehm, E.J.; Goodman, C.S.
 Cell 67, 45-57, 1991
 A:Title: Genetic analysis of growth cone guidance in Drosophila: fasciclin II function
 A:Reference number: A41054; PMID:92005695; PMID:1513618
 A:Accession: A41054
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-811 <GR2>
 A:Cross-references: UNIPROT:P34082; GB:M77165; NID:g157402; PID:g157403
 C:Genetics:
 A:Gene: FlyBase:Fa2
 A:Cross-references: FlyBase:FBgn0000635
 C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; lmr
 C:Keywords: membrane protein

Query Match 90.9%; Score 30; DB 2; Length 811;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 481 VGTGR 486

RESULT 12

B41054
 fasciclin II pr-linked splice form precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 17-Mar-2000
 C:Accession: B41054
 R:Grenningloh, G.; Rehm, E.J.; Goodman, C.S.
 Cell 67, 45-57, 1991
 A:Title: Genetic analysis of growth cone guidance in Drosophila: fasciclin II functions
 A:Reference number: A41054; PMID:92005695; PMID:1513618
 A:Accession: B41054
 A:Status: preliminary

Sat Nov 6 18:59:22 2004

us-10-054-873-3.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:20:47 Search time 4.4946 Seconds
(without alignments)
432.820 Million cell updates/sec

Title: US-10-054-873-3

Perfect score: 33
Sequence: 1 LGTGR 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	6	US-10-054-873-3	Sequence 3, Appl1
2	33	100.0	15	US-10-337-598-290	Sequence 290, App
3	33	100.0	82	US-10-437-963-168327	Sequence 168327,
4	33	100.0	107	US-10-054-873-6	Sequence 6, Appl1
5	33	100.0	135	US-10-408-7654-1525	Sequence 1525, Ap
6	33	100.0	150	US-10-054-873-7	Sequence 7, Appl1
7	33	100.0	275	US-10-369-493-12819	Sequence 12819, A
8	33	100.0	802	US-10-437-963-178776	Sequence 178776,
9	33	100.0	952	US-10-108-2604-3284	Sequence 3284, Ap
10	33	100.0	2013	US-10-282-1224-60608	Sequence 60608, A
11	33	100.0	7257	US-10-014-717-5	Sequence 5, Appl1
12	31	93.9	455	US-10-437-963-154712	Sequence 154712,
13	31	93.9	1963	US-10-437-963-155445	Sequence 155445,

14	30	90.9	24	US-09-864-761-19589	Sequence 39589, A
15	30	90.9	78	US-10-437-963-188463	Sequence 188463, A
16	30	90.9	78	US-10-767-701-56946	Sequence 56946, A
17	30	90.9	88	US-10-424-599-183036	Sequence 183036,
18	30	90.9	114	US-09-728-626-3668	Sequence 3668, Ap
19	30	90.9	186	US-10-767-701-36431	Sequence 36431, A
20	30	90.9	208	US-10-233-553-30	Sequence 30, Appl
21	30	90.9	222	US-10-240-145-161	Sequence 161, App
22	30	90.9	231	US-10-243-552-564	Sequence 564, App
23	30	90.9	268	US-10-424-599-14755	Sequence 14755,
24	30	90.9	275	US-10-424-599-168540	Sequence 168540,
25	30	90.9	275	US-10-425-114-48504	Sequence 48504, A
26	30	90.9	298	US-09-925-301-1224	Sequence 1224, Ap
27	30	90.9	298	US-10-296-115-1307	Sequence 1307, Ap
28	30	90.9	309	US-10-425-114-47615	Sequence 47615, A
29	30	90.9	309	US-10-425-114-48066	Sequence 48066, A
30	30	90.9	377	US-10-437-963-173919	Sequence 173919,
31	30	90.9	432	US-09-728-626-6081	Sequence 6081, Ap
32	30	90.9	471	US-10-029-386-31856	Sequence 31856, A
33	30	90.9	575	US-10-425-114-64569	Sequence 64569, A
34	30	90.9	575	US-10-425-114-64569	Sequence 64569, A
35	30	90.9	676	US-10-374-979-101	Sequence 101, App
36	30	90.9	676	US-10-182-9364-101	Sequence 101, App
37	30	90.9	804	US-10-128-714-3088	Sequence 3088, Ap
38	30	90.9	1028	US-10-128-714-8098	Sequence 8098, Ap
39	30	90.9	1148	US-10-437-963-14916	Sequence 14916,
40	30	90.9	1204	US-10-437-963-188462	Sequence 188462,
41	30	90.9	1233	US-10-437-963-188471	Sequence 188471,
42	30	90.9	1467	US-10-437-963-106949	Sequence 106949,
43	30	90.9	1540	US-10-437-963-188461	Sequence 188461,
44	30	90.9	2071	US-10-437-963-103806	Sequence 103806,
45	30	90.9	2978	US-10-437-963-121042	Sequence 121042,

ALIGNMENTS

RESULT 1
US-10-054-873-3
; Sequence 3, Application US/10054873
; Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chapterone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor,
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3
US-10-054-873-3

Query Match
Best Local Similarity 100.0%; Score 33; DB 13; Length 6;
Pred. No. 1,2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 1 LGTGR 6

RESULT 2
US-10-327-598-290
Sequence 280 Application US/10327598
Publication No. US20040181039A1
GENERAL INFORMATION:
APPLICANT: Krah, Eugene
APPLICANT: Guo, Hongliang
APPLICANT: Aiyappa, Ashok
APPLICANT: Lawton, Robert
TITLE OF INVENTION: Canine Immunoglobulin Variable Domains, Caninized Antibodies, and
FILE REFERENCE: 01-799-A
CURRENT APPLICATION NUMBER: US/10/327,598
CURRENT FILING DATE: 2002-12-20
PRIOR APPLICATION NUMBER: US 60/344,874
PRIOR FILING DATE: 2001-12-21
NUMBER OF SEQ ID NOS: 1139
SOFTWARE: PatentIn version 3.0
SEQ ID NO 290
LENGTH: 15
TYPE: PRT
ORGANISM: canis familiaris;
US-10-327-598-290

Query Match
Best Local Similarity 100.0%; Score 33; DB 16; Length 15;
Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 6 LGTGR 11

RESULT 3
US-10-437-963-168327
Sequence 168327 Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 168327
LENGTH: 82
TYPE: PRT
ORGANISM: Oryza sativa

FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(82)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_66852C.1.pcp
US-10-437-963-168327

Query Match
Best Local Similarity 100.0%; Score 33; DB 16; Length 82;
Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 13 LGTGR 18

RESULT 4
US-10-054-873-6
Sequence 6 Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
CLASSIFICATION: <Unknown>
FILING DATE: 22-Jan-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 107 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6
US-10-054-873-6

Query Match
Best Local Similarity 100.0%; Score 33; DB 13; Length 107;
Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 50 LGTGR 55

RESULT 5
US-10-408-765A-1525
Sequence 1525 Application US/10408765A

Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Faby, Boi D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Wernick, Gary M.
APPLICANT: Wernick, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1525
LENGTH: 135
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-1525

Query Match 100.0%; Score 33; DB 16; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 5 LGTGR 10

RESULT 6
US-10-054-873-7
Sequence 7, Application US/10054873
Publication No. US20020164712A1
GENERAL INFORMATION:
APPLICANT: Gan, Zhong Ru
TITLE OF INVENTION: Chimeric Protein Containing an
Intramolecular Chaperone-Like Sequence
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/054,873
FILING DATE: 22-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/CN98/00052
FILING DATE: 31-MAR-1998
APPLICATION NUMBER: US 09/423,100
FILING DATE: 11-DEC-2000
ATTORNEY/AGENT INFORMATION:
NAME: Mycroft, Frank J
REGISTRATION NUMBER: 46,946
REFERENCE/DOCKET NUMBER: 020167-000130US
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-054-873-7
Query Match 100.0%; Score 33; DB 13; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 93 LGTGR 98

RESULT 7
US-10-369-493-12819
Sequence 12819, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 12819
LENGTH: 275
TYPE: PRT
ORGANISM: Aspergillus nidulans
US-10-369-493-12819

Query Match 100.0%; Score 33; DB 14; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
Db 15 LGTGR 20

RESULT 8
US-10-437-963-178776
Sequence 178776, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbaruk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 178776
LENGTH: 802
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_7629C.1.pep
US-10-437-963-178776

Query Match 100.0%; Score 33; DB 16; Length 802;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGTGR 6
|||||
DB 279 LGTGR 284

RESULT 9
US-10-108-260A-3284
; Sequence 3284, Application US/10108260A
; Publication No. US2004005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US2004005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3284
; LENGTH: 952
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3284

Query Match 100.0%; Score 33; DB 15; Length 952;
Best Local Similarity 100.0%; Pred. No. 9,8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
DB 634 LGTGR 639

RESULT 10
US-10-282-122A-60608
; Sequence 60608, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60608
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60608

Query Match 100.0%; Score 33; DB 15; Length 2013;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
DB 892 LGTGR 897

RESULT 11
US-10-014-717-5
; Sequence 5, Application US/10014717
; Publication No. US20020192778A1
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/10/014,717
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US/09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-10-014-717-5

Query Match 100.0%; Score 33; DB 13; Length 7257;
Best Local Similarity 100.0%; Pred. No. 6,5e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
|||||
DB 1041 LGTGR 1046

RESULT 12
US-10-437-963-154712
; Sequence 154712, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Bouharov, Andrew A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966


```

; SEQ ID NO 154712
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(455)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_54545C.1.pep
US-10-437-963-154712

Query Match
Best Local Similarity 93.3%; Score 31; DB 16; Length 455;
Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
Db 338 MGTGR 343

RESULT 13
US-10-437-963-155445
; Sequence 155445, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155445
; LENGTH: 1963
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55207C.1.pep
US-10-437-963-155445

Query Match
Best Local Similarity 93.9%; Score 31; DB 16; Length 1963;
Pred. No. 4.5e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
Db 336 MGTGR 341

RESULT 14
US-09-864-761-39589
; Sequence 39589, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anncmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39589
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004061.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.84
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
US-09-864-761-39589

Query Match
Best Local Similarity 90.3%; Score 30; DB 9; Length 24;
Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
Db 13 LGTGR 18

RESULT 15
US-10-437-963-188463
; Sequence 188463, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21 (53221) B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ. ID NOS: 204966
SEQ. ID NO 188463
LENGTH: 78
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_85064C.1.pap
US-10-437-963-188463

Query Match 90.9%; Score 30; DB 16; Length 78;
Best Local Similarity 83.3%; Pred. No. 3.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

CY 1 LGTGPR 6
:|||||
Db 30 VGTGPR 35

Search completed: November 2, 2004, 20:59:21
Job time : 6.49446 secs


```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakematsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita S.,
RA Setine M., Ogihara T., Nishi T., Shibahara T., Tanaka K., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagai K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoki T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niimiyama K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kinoshita T., Yamashita H., Hirooka S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi F., Hara H., Hara H., Tanase T., Nomura Y.,
RA Tsujia S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki M., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Taniguchi A., Fujimura T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,
RA Matsuura T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Takasumi S., Sugano T., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Mitsuhashi S., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RA "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
RP EMBL; AK022407; BAB14030.1;
RC SEQUENCE 135 AA; 14034 MW; 0D37366C979CDDA8 CRC64;
SQ
Query Match 100.0%; Score 33; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGTGR 6
DB 5 LGTGR 10
RESULT 3
QBLUS PRELIMINARY; PRT; 277 AA.
AC QBLUS;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Bhut.
CN Name=Bhut;
OS Bordetella avium.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=521;
RV [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4169;
RX MEDLINE=21461973; PubMed=11598070;
RA Kirby A.E., Metzger D.J., Murphy E.R., Connell T.D.;
RT "Heme utilization in Bordetella avium is regulated by Bhut, a heme-
RL responsive extracytoplasmic function sigma factor."
RN Infect. Immun. 69:6951-6961(2001).
RC SEQUENCE FROM N.A.
RP STRAIN=4169;

```

```

RX MEDLINE=22215669; PubMed=12228263;
RA Murphy E.R., Sacco R.E., Dickenson A., Metzger D.J., Hu Y.,
RA Orndorff P.E., Connell T.D.;
RT "Bhut, a virulence-associated outer membrane protein of Bordetella
RT avium, is required for the acquisition of iron from heme and
RT hemoproteins."
RL Infect. Immun. 70:5390-5403(2002).
DR EMBL; AY095952; AAM28270.1;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2.1.
SQ SEQUENCE 277 AA; 28938 MW; F9CDDCCDDA37B4D CRC64;
Query Match 100.0%; Score 33; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1,8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGTGR 6
DB 252 LGTGR 257
RESULT 4
Q7W361 PRELIMINARY; PRT; 287 AA.
AC Q7W361;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative heme binding protein.
GN Name=Bhut; OrderedLocustNames=BPP4187;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
CX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bacon N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jags K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitch B., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:332-40(2003).
DR EMBL; EX640436; CAE39466.1;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2.1.
KV Complete proteome.
SQ SEQUENCE 287 AA; 29393 MW; 86F8317AD5241C40 CRC64;
Query Match 100.0%; Score 33; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1,8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LGTGR 6
DB 259 LGTGR 264
RESULT 5
Q7WEH9 PRELIMINARY; PRT; 287 AA.
Q7WEH6

```

```

AC Q7MEH8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative hemin binding protein.
GN Name=hnt; OrderedLocustNames=BB4657;
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OK NCBI_TaxID=518;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parthill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bacon N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtz S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 29363 MW; 96F9317AC5251031 CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 287;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
DB 259 LGTGR 264

RESULT 6
ID Q7VSO6; PRELIMINARY; PRT; 269 AA.
AC Q7VSO6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative hemin binding protein.
GN Name=hnt; OrderedLocustNames=BP0345;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OK NCBI_TaxID=520;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parthill J., Sebailia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bacon N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtz S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";

```

```

RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640412; CA64677.1;
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0006827; P:high affinity iron ion transport; IEA.
DR InterPro; IPR002491; Peripla_BP.
DR Pfam; PF01497; Peripla_BP_2; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 29505 MW; 3B80C28C1D8940AD CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 289;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
DB 261 LGTGR 266

RESULT 7
ID Q8P623; PRELIMINARY; PRT; 491 AA.
AC Q8P623;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein-glutamate methyltransferase.
GN Name=cheB; OrderedLocustNames=XCC2822;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OK NCBI_TaxID=340;
RN 111
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Queiroz R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., de Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canavari F., Cardoso J., Chamberg F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formigueri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Melandris J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Szejnolt L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Tindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA "Secubal J.C., Kitchima J.P.";
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AEO12394; AAM42094.1;
DR HSPD; P04432; ICHD.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000673; CheB_methylase.
DR InterPro; IPR01247; Chmx_methylase.
DR Pfam; PF01339; CheB_methylase; 1.
DR PIRSF; PIRSF036461; Chmx_methylase; 1.
DR ProDom; PD003328; CheB_methylase; 1.
DR PROSITE; PS0123; CheB_1; MHC.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 491 AA; 51780 MW; 379E3413A027F619 CRC64;

Query Match
Best Local Similarity 100.0%; Score 33; DB 2; Length 491;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Cy 1 LGTGR 6
DB 248 LGTGR 253

RESULT 6
ID Q810F1 PRELIMINARY; PRT; 563 AA.
AC Q810F1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 05-JUN-2004 (TREMBlrel. 27, Last sequence update)
DE C31338-BA (A027833p).
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Ephydroidea; Drosophilidae; Diptera; Brachycera; Muscomorpha;
OC NCBI_TaxID=7227;
RN SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazef R.G., Champe M., Pfeiffer B.D.,
RA Wan X.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Buzam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport I.B., Davies P.,
RA Chery P.B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Doming M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evansglista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glöckner A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.U., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegawa C.,
RA Jaitani M., Kalush F., Kappen G.H., Ke Z., Kermanson J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA LaRoche J., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese W.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri U.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Facel S., Adams W., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirska R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weisenbach G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RA melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]

RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirska R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RL a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,
RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo B., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
RN SEQUENCE FROM N.A.
RP FLYBASE;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RP FLYBASE;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Paffan D., Frise E.,
RA George R., Gonzalez M., Guarin B., Krommiller B., Li P., Lao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AEO03603; AAN13275.1;
DR EMBL: BT001356; AAN71111.1;
DR FLYBASE; FBN0051538; CG31538.
SQ SEQUENCE 563 AA; 63800 MW; 8E054274E710C563 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 3; +e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 LGTGR 6
DB 269 LGTGR 274

RESULT 9
ID Q8NEG7 PRELIMINARY; PRT; 585 AA.
AC Q8NEG7;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Similar to mouse 1700027J05R1k protein.
GN Name=MG033592;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=Testis;
RC MEDLINE=22388257; PubMed=12477932;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausberg R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

```

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant J.L., Scheetz T.E.,
 RA Brownstein W.J., Usdin T.B., Toshiyuki S., Caranici P., Pange C.,
 RA Raha S.S., Lottellano N.A., Peters G.J., Abramson R.D., Millamy S.D.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiy S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalski U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Warr M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Query Match 100.0%; Score 33; DB 2; Length 585;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 5 LGTGR 10

RESULT 10
 O6YBM4 PRELIMINARY; PRT; 736 AA.

AC O6YBM4;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE TACC3.
 OS Name=TACC3;
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA PubMed=15207008;
 RA Still I.H., Veltikornumakanau A.K., DiMatteo A., Liang P.,
 RT "Structure-function evolution of the transforming acidic coiled coil
 RT genes revealed by analysis of phylogenetically diverse organisms.";
 RL BMC Evol. Biol. 4:16-16(2004).
 DR EMBL; AY161270; AA025635.1; -
 SO SEQUENCE 736 AA; 77061 MW; A798FB1C17EF3C8 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 736;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 207 LGTGR 212

RESULT 11
 AA025635 PRELIMINARY; PRT; 736 AA.
 AC AA025635;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE TACC3.
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

GN TACC3
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Still I.H., Liang P.,
 RT "Evolutionary analysis reveals lack of evidence for a fourth TACC gene
 RT member in vertebrates.";
 RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY161270; AA025635.1; -
 SO SEQUENCE 736 AA; 77061 MW; A798FB1C17EF3C8 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 736;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 207 LGTGR 212

RESULT 12
 O6ZRS4 PRELIMINARY; PRT; 896 AA.

AC O6ZRS4;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Hypothetical protein FLJ46145.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE=Testis;
 RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,
 RA Matsunoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,
 RA Matsunuma Y., Moriya S., Chiba S., Momiyama H., Onogawa S.,
 RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Terashima Y.,
 RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,
 RA Watanabe M., Sugiyama T., Irie Y., Otsuki T., Sato H., Wakamatsu A.,
 RA Ishii S., Yamamoto J., Isoro Y., Kawai-Hiro Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Wakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuho Y., Nagai K., Isegaki T.,
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK128025; BAC87235.1; -
 SO SEQUENCE 896 AA; 98946 MW; DEC9FE06CF7B2C0 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 896;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LGTGR 6
 DB 578 LGTGR 583

RESULT 13
 BAC87235 PRELIMINARY; PRT; 896 AA.
 AC BAC87235;
 DT 02-MAR-2004 (TREMBlrel. 27, Created)
 DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)
 DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)
 DE CDNA FLJ46145 f1s, clone TEST14000370.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI_TaxID=96606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohara N.,
 RA Matsumoto K., Hirono M., Sano S., Nomura R., Yoshikawa Y.,
 RA Matsumura Y., Moriya S., Chida E., Momiyama H., Onogawa S.,
 RA Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,
 RA Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Tetsuhara Y.,
 RA Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A.,
 RA Ichii S., Yamamoto J., Iseno Y., Kawai-Hito Y., Saito K., Nishikawa T.,
 RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
 RA Kikuchi H., Kanda K., Magatsuma M., Murakawa K., Kanehori K.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahari K., Masuko Y., Nagai K., Isogai T.,
 RT "NEBO human cDNA sequencing project."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK128026; BAC87235.1;
 SO SEQUENCE 996 AA; 98946 MW; DBC9EFOB6CF732C0 CRC64;
 Query Match 100.0%; Score 33; DB 2; Length 996;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 DB 578 LGTGR 583
 RESULT 14
 QP5K6 PRELIMINARY; PRT; 946 AA.
 ID QP5K6;
 AC QP5K6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Related to glucan 1, 4-alpha-glucosidase.
 CN Name=EP32L21.230;
 OS Neurospora crassa;
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schulte U., Algen V., Hohnselt J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Neues H.W., Mannhaupt G.,
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA German Neurospora genome project;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL356172; CAB91691.2;
 DR PIR: T49710; T49710.
 SO SEQUENCE 946 AA; 101461 MW; A8564328338B6B1C CRC64;
 Query Match 100.0%; Score 33; DB 2; Length 946;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 DB 807 LGTGR 812
 RESULT 15
 Q8KLSO PRELIMINARY; PRT; 1095 AA.
 ID Q8KLSO;
 AC Q8KLSO;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Putative histidine protein kinase.
 GN Name=chea3;

OS Rhodospirillum rubrum (Rhodospirillum rubrum)
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Rhodospirillum.
 OX NCBI_TaxID=1063;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MS81;
 RA Porter S.L., Warren A.V., Armistead J.P.,
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ488585; CAD32761.1;
 DR GO: GO:0005622; C:intracellular; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0016301; F:kinase activity; IEA.
 DR GO: GO:0000155; F:two-component sensor molecule activity; IEA.
 DR GO: GO:0006935; P:chemotaxis; IEA.
 DR GO: GO:0000160; P:two-component signal transduction system (p. . .); IEA.
 DR InterPro: IPR003594; ATPbind_ATPase.
 DR InterPro: IPR008207; HPT.
 DR InterPro: IPR008208; HPT_N.
 DR Pfam: PF01584; CheW; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR Pfam: PF01627; HPT; 1.
 DR ProDom: PD003142; HPT_N; 1.
 DR SMART: SM00260; CheW; 1.
 DR SMART: SM00073; HPT; 1.
 DR ProSite: PS50894; HPT; 1.
 KM kinase
 SO SEQUENCE 1095 AA; 114521 MW; FA3CF5A0EBAF3F0E CRC64;
 Query Match 100.0%; Score 33; DB 2; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 6.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LGTGR 6
 DB 1031 LGTGR 1036

Search completed: November 2, 2004, 20:20:31
 Job time: 7.8181 secs